

Click on 'Bookmarks' in the Left-Hand menu and then click on the required section - answers to Q questions at end of section.

8.1a z-Area Calculations for a Normal Distribution

8.1a.1 Standard Normal Distribution

The **Standard Normal Distribution, $N(0,1)$** , is a specific normal distribution, which has a *mean value* of **0.0**, a *standard deviation* of **1.0** - see Figure 8.1a.1.

The variable on the abscissa (horizontal axis) of the standard normal distribution is usually written as z .

The probabilities of recording a value that falls between specific z -values can be calculated from standard Tables - Appendix II.

In particular, Figure 8.1a.1 shows the probabilities (areas) of recording values between *integer* z -values.

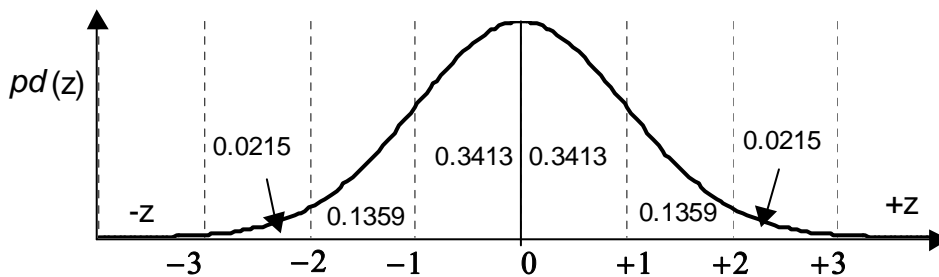


Figure 8.1a.1 Some Probability Areas for a Standard Normal Distribution

The probability of recording a value between z_1 and z_2 can be written as $p(z_1 < z < z_2)$.

The *probability*, $p(z_1 < z < z_2)$, is equal to the *area* under the standard normal distribution between values z_1 and z_2 .

The *total probability (area)* for *all* values of z , $p(-\infty < z < +\infty) = 1.0$, and the probability area for each 'half' of the distribution, $p(-\infty < z < 0) = p(0 < z < +\infty) = 0.5$.

The normal distribution is symmetrical - the areas on the negative side of the distribution are the same as the equivalent areas on the positive side.

Example 8.1a.1

Using the probability areas in Figure 8.1a.1 we can calculate the probabilities of recording z values within specific ranges:

(it is often necessary to add or subtract areas to get the desired final area)

i)	$p(0 < z < +1) = 0.3413 \approx 34\%$
ii)	$p(-1 < z < +1) = 0.3413 + 0.3413 = 0.6826 \approx 68\%$
iii)	$p(+1 < z < +2) = 0.1359$
iv)	$p(0 < z < +2) = p(0 < z < +1) + p(+1 < z < +2) = 0.3413 + 0.1359 = 0.4772 \approx$

	47.7%
v)	$p(-2 < z < +2) = 2 \times 0.4772 = 0.9544 \approx 95\%$
vi)	$p(-\infty < z < 0) = 0.5 = 50\%$
vii)	$p(-\infty < z < 2) = p(-\infty < z < 0) + p(0 < z < +2) = 0.5 + 0.4772 = 0.9772$
viii)	$p(2 < z < \infty) = 1.0 - p(-\infty < z < +2) = 1.0 - 0.9772 = 0.0228 \approx 2.3\%$ Note in this case, how the area in the right hand 'tail' ($z > 2$) is calculated by taking the area between $z = -\infty$ and $z = 2$ away from the total area (1.0).

Q8.1a.1 Use the data in Figure 8.1a.1 to calculate the following probability areas for the standard normal distribution, $N(0,1)$:

- i) $p(+2 < z < +3)$
- ii) $p(0 < z < +3)$
- iii) $p(-3 < z < +3)$
- iv) $p(+3 < z < \infty)$
- v) $p(-\infty < z < +3)$

8.1a.2 Calculating z Probability Areas

The fractional areas under the *standard normal distribution* are commonly published as a function of the *abscissa variable*, which is usually written as **z**.

The **z-Area Table** (Appendix II) gives the total probability (= area), of readings between $z = -\infty$ (the extreme left of the distribution) and the selected value of **z**.

For example, the probability area, $A_2 (= 0.7517)$, shown in Figure 8.1a.2(a) gives the probability of recording a value between $z = -\infty$ and $z_2 = +0.68$:

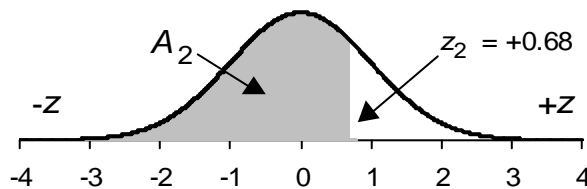


Figure 8.1a.2(a) Probability Area, $A_2 = 0.7517$

and the probability area, $A_1 (= 0.0735)$ in Figure 8.1a.2(b) gives the probability of recording a value between $z = -\infty$ and $z_1 = -1.45$:

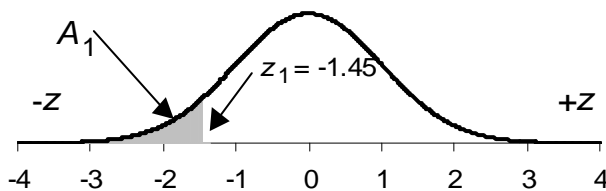


Figure 8.1a.2(b) Probability Area, $A_1 = 0.0735$

By taking the *difference* between the areas in 8.1a.2(a) and 8.1a.2(b), we can then calculate the probability area, $A = A_2 - A_1$ ($= 0.6782$) in Figure 8.1a.2(c), which gives the probability of recording a value between $z_1 = -1.45$ and $z_2 = 0.68$: 0.68:

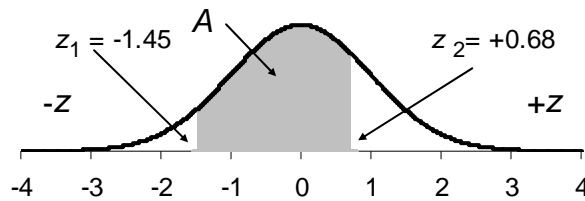


Figure 8.1a.2(c) Probability Area $A = A_2 - A_1$

To calculate the probability (area) between any two specific values of z :
(e.g. from $z_1 = -1.45$ to $z_2 = +0.68$):

1. Look up the areas, A_1 and A_2 , from the z -Area Table for z_1 and z_2 .
e.g. for $z_2 = 0.68$, the area, $A_2 = p(-\infty < z < z_2) = 0.7517$ - see Figure 8.1a.2(a)
2. If the z value is **negative**, look up the area for the equivalent *positive* z value, and subtract the area from 1.0000.
e.g. for $z_1 = -1.45$, look up the probability area for $z = +1.45$, $p(-\infty < z < +1.45)$, which gives 0.9265, and then the area, $A_1 = p(-\infty < z < -1.45) = 1.0000 - 0.9265 = 0.0735$ - see Figure 8.1a.2(b)
3. Calculate the difference between the two values, $A = A_2 - A_1$
 $A = A_2 - A_1 = 0.7517 - 0.0735 = 0.6782$ - see Figure 8.1a.2(c)
4. The probability of randomly selecting a value between $z = -1.45$ and $z = +0.68$ from the standard normal distribution is given by the probability area, A , shown in Figure 8.1a.2(c). The probability, A , = 0.6782

It is useful (as in Figure 8.1a.2) to draw small sketches of the areas being calculated from the z -Area Table, and use common sense to check correct addition or subtraction.

Areas in the positive (or upper) 'tail' of the distribution can be calculated by *subtracting the z -area from 1.0*.

For example, the probability area for $z > 2.0$ can be written as $p(2.0 < z)$, and is calculated as follows:

$$p(2.0 < z) = p(2.0 < z < +\infty) = 1.0000 - p(-\infty < z < 2.0) = 1.0 - 0.9772 = 0.0228.$$

Areas in the lower tail are equal (symmetrical) to the equivalent upper tail value, i.e. $p(z < -2.0) = p(2.0 < z) = 0.0228$.

The probability areas for the standard normal distribution can also be calculated using the functions NORMSDIST or NORMDIST in EXCEL.

Example 8.1a.2

Use statistical Tables to calculate the following probability areas for the standard normal distribution, $N(0,1)$:

(it may be helpful to draw small sketches of the areas on a normal distribution)

- i) $p(-\infty < z < 0) = 0.5000$
(exactly half of the distribution)
- ii) $p(-\infty < z < +0.9) = 0.8159$
(value directly from Tables)
- iii) $p(0 < z < +0.9) = p(-\infty < z < +0.9) - p(-\infty < z < 0) = 0.8159 - 0.5000 = 0.3159$
(difference between (ii) and (i))
- iv) $p(-0.9 < z < 0) = 0.3159$
(area is symmetrical with (iii))
- v) $p(-0.9 < z < +0.9) = 0.3159 + 0.3159 = 0.6318$
(addition of (iii) and (iv))
- vi) $p(+0.9 < z < +\infty) = 1.0000 - p(-\infty < z < +0.9) = 1.0000 - 0.8159 = 0.1841$
(gives the upper tail area to the right of $z = 0.9$)
- vii) $p(-\infty < z < -0.9) = 1.0000 - p(-\infty < z < +0.9) = 0.1841$
(the lower tail area is found by calculating the equivalent upper tail area)
- viii) $p(-\infty < z < -1.62) = 1.0000 - p(-\infty < z < +1.62) = 1.0000 - 0.9474 = 0.0526$
(taking value from Tables)
- ix) $p(-1.62 < z < +0.9) = p(-\infty < z < +0.9) - p(-\infty < z < -1.62) = 0.8159 - 0.0526 = 0.7633$
(difference between (ii) and (viii))

Q8.1a.2 Use statistical Tables to calculate the following probability areas for the standard normal distribution, $N(0,1)$:

- i) $p(-\infty < z < +0.75)$
- ii) $p(0 < z < +0.75)$
- iii) $p(-0.75 < z < 0)$
- iv) $p(-0.22 < z < 1.31)$
- v) $p(1.31 < z < 2.0)$
- vi) $p(-1.56 < z < -0.42)$
- vii) $p(2.5 < z) = p(2.5 < z < +\infty)$
- viii) $p(z < -1.78) = p(-\infty < z < -1.78)$

8.1a.3 Reverse z Calculations

There are occasions when it is necessary to carry out a z calculation *backwards*, i.e. starting with a probability area, A , and then looking up the equivalent z -value from the Tables.

It is possible to calculate z -values from probabilities for the standard normal distribution by using the function NORMSINV or NORMINV in EXCEL.

Example 8.1a.3

Find the z-value such that the probability of getting a value greater than this value is 5%.

In this case, we need to find a value, z' , such that $p(z' < z) = 0.0500$.

If $p(z' < z) = p(z' < z < \infty) = 0.0500$, we can deduce that $p(-\infty < z < z') = 0.9500$.

Hence we need to look in the z-Area Table to find a value of z' such that the probability area, $A = 0.9500$.

We find that $z = 1.64$ gives $A = 0.9495$ and $z = 1.65$ gives $A = 0.9505$.

By interpolation, we infer that the value of z' must lie midway between these values, giving $z' = 1.645$

Q8.1a.3

Use statistical Tables to calculate the following values of z' for the standard normal distribution, $N(0,1)$:

(It would be useful to draw the probability areas on small sketch graphs - this will help work out the additions and subtractions required)

i) Value of z' , such that $p(0 < z < z') = 0.3944$

Hint: $p(0 < z < z') = p(-\infty < z < z') - 0.5000$

ii) Value of z' , such that $p(z' < z < 0) = 0.3944$

Hint: By symmetry $p(z' < z < 0) = p(0 < z < z')$ where $z' = -z''$, then use result from (i).

iii) Value of z' , such that $p(z' < z) = p(z' < z < +\infty) = 0.025$

Hint: $p(z' < z < +\infty) = 1.0000 - p(-\infty < z < z')$

iv) Value of z' , such that $p(z' < z) = p(z' < z < +\infty) = 0.005$

Hint: As for (iii)

v) Value of z' , such that $p(z < z') = p(-\infty < z < z') = 0.025$

Hint: By symmetry $p(z < z') = p(z' < z)$ where $z' = -z''$, then use result from (iii).

8.1a.4 Common Probability Areas

We can use the probability areas from the standard normal distribution to work out some common probabilities associated with *all* normal distributions:

- 68.3% ($\approx 2/3$) of all data points lie within \pm ONE standard deviation from the mean.
- 95.4% ($\approx 95\%$) of all data points lie within \pm TWO standard deviations from the mean.
- 99.7% of all data points lie within \pm THREE standard deviations from the mean.
- 90% of all data points lie within ± 1.64 standard deviations from the mean.
- 95% of all data points lie within ± 1.96 ($\approx \pm 2$) standard deviations from the mean.
- 99% of all data points lie within ± 2.58 standard deviations from the mean.

Q8.1a.4 Use statistical Tables (or the EXCEL function NORMSINV), to calculate the following values of z' in the standard normal distribution such that

- i) 2.5% of the probability area lies in the region defined by: $z' < z$
- ii) 0.5% of the probability area lies in the region defined by: $z' < z$
- iii) 95% of the probability area lies in the region defined by: $-z' < z < z'$
- iv) 99% of the probability area lies in the region defined by: $-z' < z < z'$

8.1a.5 Using z-Values for ANY Normal Distribution

The probability areas (between x_1 and x_2) for *any* normal distribution (with mean, μ , and standard deviation, σ) can be calculated by comparing it with the *identical* probability areas between equivalent points (z_1 and z_2) of the standard normal distribution.

For example, we will compare the probability area for the distribution $N(80,4)$ in Figure 8.1a.3(a) with the same probability area in the standard normal distribution, $N(0,1)$ in Figure 8.1a.3(b).

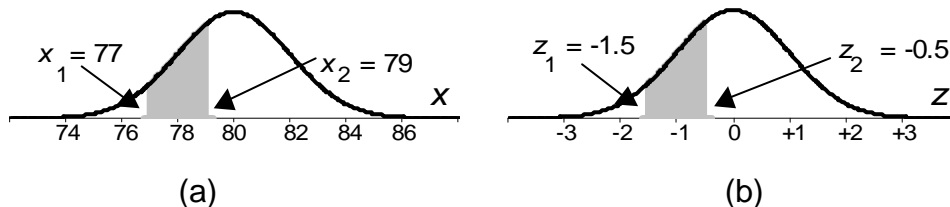


Figure 8.1a.3 Identical Probability Areas in x and z distributions.

We can see from Figure 8.1a.3 that the probability area between $x_1 = 77$ and $x_2 = 79$ in the normal distribution, $N(80,4)$, **is the same** as the probability area between $z_1 = -1.5$ and $z_2 = -0.5$ in the standard normal distribution, $N(0,1)$.

The 'z-value', z_i , in Figure 8.1a.3(b) that is equivalent to the 'x-value', x_i , in Figure 8.1a.3(a) is given by the important equation:

$$z_i = \frac{(x_i - \mu)}{\sigma} \quad [8.1a.1]$$

We outline, in example 8.1a.4 below, the general procedure for addressing this type of problem.

Example 8.1a.4

Using the distribution in Figure 8.1a.3 with a normal distribution of data values with mean = 80.0 and standard deviation = 2.0, calculate the probability of recording a data value between 77.0 and 79.0.

1. Calculate the z_i values equivalent to the particular x_i values using [8.1a.1] with $\mu = 80$ and $\sigma = 2$:

$$z_i = \frac{(x_i - 80)}{2}$$

For $x_1 = 77$, the equivalent $z_1 = (77 - 80)/2 = -1.5$

For $x_2 = 79$, the equivalent $z_2 = (79 - 80)/2 = -0.5$

2. Calculate the probability areas from the z-values
 Probability area for $z_1 = -1.5$, $A_1 = 0.0668$
 Probability area for $z_2 = -0.5$, $A_2 = 0.3085$
3. Calculate the area between the two values
 $A = A_2 - A_1 = 0.3085 - 0.0668 = 0.2417$
 This is the probability of recording a data value between -1.5 and -0.5 in the standard normal distribution.
4. The probability areas in Figures 8.6(a) and (b) are the same. Hence, we can also say that there is a probability of 0.2417 (= 24.17%) of recording a data value between 77.00 and 79.00 for a normal distribution of data values with mean = 80.0 and standard deviation = 2.0
 Answer = 0.2417

- Q8.1a.5** A normal probability distribution along the variable axis x is described by $N(5,9)$. Calculate
- i) mean value of the distribution
 - ii) standard deviation of the distribution
 - iii) z-value equivalent to the mean value, $x = 5$
 - iv) z-value equivalent to $x = 8$
 - v) z-value equivalent to $x = 2$
 - vi) z-value equivalent to $x = 11$
 - vii) z-value equivalent to $x = -1$
 - viii) probability area, $p(5 < x < 11)$, between the mean value, $x = 5$, and $x = 11$
 - ix) probability area, $p(2 < x < 8)$, between $x = 2$ and $x = 8$
 - x) probability area, $p(-1 < x < 11)$, between $x = -1$ and $x = 11$
 - xi) probability area, $p(11 < x < \infty)$, greater than $x = 11$

The cumulative probability areas for *any* normal distribution can also be calculated using the function NORMDIST in EXCEL.

Example 8.1a.5

The distribution of (integer) marks for a particular examination is described by a normal distribution with

Mean Value, $\bar{x} = 53.50$

Standard Deviation, $s = 11.47$

For a group of 200 students, use EXCEL to estimate the numbers of students falling into each of the ranges given in the Table below.

- Take the boundaries between each range at values 9.5, 19.5, etc.

- Use the function NORMDIST to calculate the *cumulative probability* up to each range boundary, e.g. NORMDIST(39.5,53.5,11.47,TRUE) calculates the cumulative probability up to 39.5.
- Calculate the *probability for each range* by taking the difference between successive values for the cumulative probability.
- Multiply the probability for each range by 200.
- The results are given in the Table below (with probabilities to 3 decimal places and the final estimated numbers given to 2 decimal places):

Range	0-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90 +
Boundary	9.5	19.5	29.5	39.5	49.5	59.5	69.5	79.5	89.5	100
Cumulative Prob.	0.000	0.001	0.018	0.111	0.364	0.700	0.918	0.988	0.999	1.000
Probability	0.000	0.001	0.017	0.093	0.253	0.336	0.219	0.070	0.011	0.001
Number	0.01	0.29	3.34	18.58	50.50	67.18	43.79	13.96	2.17	0.16

8.1a.6 Reverse Calculations for ANY Normal Distribution

When performing 'reverse' calculations (i.e. finding values of x from z values) it is necessary to calculate x_i from z_i , and this requires a rearrangement of [8.1a.1]:

$$x_i = \mu + (z_i \times \sigma) \quad [8.1a.2]$$

Q8.1a.6 A normal probability distribution along the variable axis x is described by $N(5,9)$. Calculate the following:

- i) mean value and the standard deviation of the distribution
- ii) x values corresponding to one standard deviation from the mean value
- iii) number of 'standard deviations' that the point $x = 2$ is from the mean value
- iv) z -value corresponding to $x = 2$
- v) value of x such that the probability of getting a value with x greater than this value is equal to 5%
- vi) symmetrical range about the mean, within which there is a 90% probability of selecting a value from the distribution
- vii) symmetrical range about the mean, within which there is a 95% probability of selecting a value from the distribution

8.1a.7 Answers to Q questions

- Q8.1a.1 0.0215, 0.4987, 0.9974, 0.0013, 0.9987
 Q8.1a.2 0.7734, 0.2734, 0.2734, 0.4920, 0.0723, 0.2778, 0.0062, 0.0375
 Q8.1a.3 1.25, -1.25, 1.96, 2.575, -1.96
 Q8.1a.4 1.96, 2.58, 1.96, 2.58
 Q8.1a.5 5, 3, 0, 1, -1, 2, -2, 0.4772, 0.6826, 0.9544, 0.0228
 Q8.1a.6 (i) 5,3 (ii) 2,8 (iii) -1 (iv) -1 (v) 9.935 (vi) 0.065 to 9.935
 (vii) -0.88 to 10.88